At section 2 of the office action, claims 1, 3-48 are rejected under 35 U.S.C. 112, second paragraph, as being indefinite for failing to particularly point out and distinctly claim the subject matter which applicant regards as the invention. The Examiner states that claims 1, 19, 22, 27, 31 and 32 have the limitation of segmenting audio signals based upon audio characteristics, but it is not clear as to which segmenting aspect of the disclosure this refers.

The Examiner further states that the specification discloses two aspects of segmenting: 1) the sub-block 12 in Figure 4, based on the input speech signal 110, generates segmented audio with associated parameters 112, and 2) the sub-block 20 segments the audio signal based on degree of voicing, etc.

The Examiner errs in stating that the sub-block 12 in Figure 4 generates segmented audio signal with associated parameters 112 based on the input speech signal 110.

It is respectfully submitted that the sub-block 12 in Figure 4 is a parameter extraction unit which is only used to extract unquantized parameters from the input speech signal 110 and provides the extracted parameters 112 to the compression module 20 (Figure 4; p.13, lines 8-13). In a typical parametric speech coder, the extracted parameters include linear prediction coefficients, speech energy (gain), pitch and voicing information (p.11, lines 24-25). Based on the behavior of the parameters, the compression module 20 carries out the segmentation of the input speech signal (p. 13, lines 21-24). An example of segmentation is shown in Figures 3a-3d, wherein the vertical dashed lines are segments boundaries. Segmentation is based on voicing and gain parameters (p.12, line 29-p.13, line 1).

Segmentation means sectioning, partitioning or dividing. There is no indication in the disclosure that the parameter extraction unit 12 partitions or divides the input speech signal into separated segments through the parameter extraction process, even though parameter extraction can be carried out in regular intervals (p.11, lines 26-32).

Thus, in the disclosure, only one block 20 in Figure 4 is used for audio signal segmentation.

For the above reasons, the 112 rejection should be withdrawn.

At section 3, claims 1, 3-14, 19-21, 26-37, 39-44 and 46-48 are rejected under 102(b) as being anticipated by *Gersho et al.* (U.S. Patent No. 6,311,154, hereafter referred to as *Gersho*). The Examiner states that *Gersho* discloses segmenting {partitioning or classifying} the audio

input signal into a plurality of segments {frames} based on the audio characteristics {classes} of the audio signal. The Examiner points to col.4, lines 25-27 to show that *Gersho* discloses classifying the frames in the speech signal into one of the plurality of classes.

The Examiner errs in two aspects:

1) classifying is not the same as segmenting or partitioning; and

2) *Gersho* classifies each of the frames into classes only after segmenting or partitioning the speech signal into frames (col.4, lines 23-34).

According to *Gersho*, for the purpose of performing linear-predictive (LP) analysis on the input speech, and for the purpose of packaging the data to be transmitted into a fixed number of bits for each fixed frame interval, the speech encoder has a fixed (basic) frame structure. Each basic frame is partitioned or segmented into M equal or nearly equal length basic subframes (col. 7, lines 18-26; Figure 2). According to *Gersho*, in conventional analysis-by-synthesis (AbS) coding schemes, the excitation signal for each subframe is selected by a search operation. It is difficult or impossible to obtain an adequately precise representation of the excitation segment using the conventional schemes (p.7, lines 27-33).

Gersho sets out to improve the AbS coding method by locating the actual time location of the active intervals in a sub-frame so that the coding effort can be concentrated with the windows corresponding to the active intervals. Active intervals are certain naturally-occurring intervals of the excitation signal which contain most of the important activity (col.7, lines 34-50). *Gersho* adaptively modifies the sub-frame boundaries and determines the window sizes and locations within sub-frames (col.2, lines 46-50). *Gersho* uses a pattern classifier to determine a classification that best describes the character of the speech signal in each frame (col.2, line 56-64). The method for coding a speech signal, according to *Gersho*, includes: 1) partitioning samples of a speech signal into frames; 2) deriving a residual signal for each frame; 3) classifying the speech signal in each frame into one of a plurality of classes; 4) identifying the location of at least one window in the frame by examining the residual signal for the frames; and 5) encoding the excitation for the frame based on the class of the frame (col.4, lines 23-34).

According to *Gersho*, class information is <u>not</u> available before the speech signal is segmented or partitioned into frames. Even in the conventional AbS schemes, excitation is searched after the speech signal is segmented into frames and into sub-frames. *Gersho* does not disclose segmenting the input speech signal into segments <u>based on</u> classes.

In contrast, the claimed invention is concerned with segmenting (partitioning) an audio signal into a plurality of segments based on audio characteristics of the audio signal, the audio characteristics indicative of parameters in a parametric representation of the audio signal, wherein the characteristics include voicing characteristics, energy characteristics, pitch characteristics in the segments of the audio signal.

Gersho does not disclose segmenting the input speech signal into segments based on the audio characteristics of the audio signal.

For the above reasons, *Gersho* fails to anticipate claims 1, 3-14, 19-21, 26-37, 39-44 and 46-48.

At section 5, claims 15-18, 22-25, 38 and 45 are rejected under 102(e) as being anticipated by *Sinha et al.* (U.S. Patent No. 7,191,136 B2, hereafter referred to as *Sinha*). In rejecting those claims, the Examiner states that *Sinha* discloses segmenting the audio signal into a plurality of segments based on audio characteristics of the audio signal (by high pass filtering the input audio signal (col. 4, lines 47-51) and then performing a non-linear parametric representation of the signal (col. 4, lines 53-59)).

It is respectfully submitted that claims 15-18 are dependent from claim 1 which includes the limitation of:

segmenting an audio signal into a plurality of segments based on audio characteristics of the audio signal, the audio characteristics indicative of <u>parameters</u> in a parametric representation of the audio signal.

Claims 22-25, 38 include the limitation of

an adjustment module for adjusting one or more <u>parameters</u> based on the audio characteristics for providing an adjusted representation of the parameters, wherein said adjusting

comprises segmenting the audio signal into a plurality of segments based on the characteristics of the audio signals.

Claim 45 is dependent from claim 19 which includes the limitations of:

an input for receiving audio data indicative of a plurality of parameters in an adjusted representation, wherein the audio data comprises a plurality of segments indicative of an input audio signal having audio characteristics and wherein the segments are obtained based on the audio characteristics and encoded with a plurality of encoding settings based on the audio characteristics; and

a module, responsive to the audio data, for generating a further audio signal based on the adjusted representation and the encoding settings.

Thus, claims 15-18, 22-25, 38 and 45 include the limitation that the input audio signal is segmented based on audio characteristics indicative of parameters in a parametric representation.

Sinha is concerned with a coding scheme which compresses information consisting of coded low frequency components as well as parametric representations for the high frequency components from the high pass filter (Abstract, column 4, lines 44-49). In particular, *Sinha* allows the input signal to pass through both a high pass filter and a low-pass filter so that the audio components in the high-frequency range and the audio components in the low-frequency range are encoded using different models. While the audio components can be encoded with parameters in a parametric representation and the audio characteristics of audio components can be indicative of parameters in the parametric representation, high frequency range or low frequency range is not a parameter in the parametric representations. Parameters, such as linear prediction coefficients, speech energy (gain), pitch and voicing information, can be used for audio signal synthesis. *Sinha* does not disclose or suggest that the input audio signal is segmented based on audio characteristics indicative of parameters in a parametric representation.

For the above reasons, *Sinha* fails to anticipate claims 15-18, 22-25, 38 and 45.

CONCLUSION

Claims 1 and 3-48 are allowable. Early allowance of all pending claims is earnestly solicited.

Respectfully submitted,

Ken Lao

Kenneth Q. Lao
Registration No. 40,061

Date: Nov. 6, 2008

WARE, FRESSOLA, VAN DER SLUYS
& ADOLPHSON LLP
Bradford Green, Building 5
755 Main Street, PO Box 224
Monroe, CT 06468
(203) 261-1234